(* COUNTS *)

30x

Ordering factors by relationship

		Exh.	Page
Introduction	30x-1	A	
		B	
30xA. Two unordered responses	30x-1	1	30x-3/4
- bundles of 2-by-2 tables	30x-1	2	30x-5
		3	30x-8/9
30xB. Untwisting signs	30x-2	C	
- a weaker example -- Woo	30x-7	4	30x-11
		5	30x-13
30xC. "All cuts" approaches	30x-10	6	30x-14
- back to the Badeners	30x-10	7	30x-17/21
- unrestricted or restricted?	30x-15	H	
		P	
30xD. A restricted example	30x-15		
- "alternating" goes where?	30x-15		
**********			
30xH. How far have we come?	30x-23		
**********			
30xP. Problems for chapter 30x	30x-24		

30x

Ordering factors by relationship

We have learned how to do many more interesting things with -- and learn much more from -- two way tables of counts involving factors whose versions are ordered. We ought, then, to be interested in a variety of ways of finding order among versions. We saw one approach, briefly, toward the end of section 30F. There order in one factor was used to induce order in another -- a relatively straightforward approach.

We are now ready to try rather more subtle approaches -- approaches in which both factors have their versions ordered on the basis of their mutual relationships. To do this means to use the counts themselves to order the factors that describe them.

30xA. Two unordered responses

Responses whose versions are unordered cry out to have their versions ordered. It may not be feasible -- or even possible -- to answer a particular cry of this sort, but we should be ready to try. Used properly, ff-logs can help us in this.

* bundles of 2-by-2 tables *

A first step toward doing this is to treat our two-way table as a bundle of 2-by-2 tables, a possibility mentioned in the introduction to chapter 30. The natural way for us to

attack a bundle of 2-by-2 tables is to calculate a bundle of ff-logs. Exhibit 1 applies this technique to the hair and eye color data of exhibits 2 to 8 of chapter 30.

As those who do the problems of panel E of this exhibit will have learned, the residuals obtained from an elementary analysis of the (two-way) bundle of 2-by-2 tables seem to be trying to tell us about much the same things as the residuals from the analyses of log counts we made in the last chapter (section 30C). Unless the new approach shows an obvious advantage, we are likely to stick to the earlier one, since we need much more arithmetic to do the ff-log analysis.

30xB. Untwisting signs

Working with 2-by-2 tables clearly allows us to do something new, however -- something that reflects a simple truth about ordered versions. Putting each single version of a response on the "same side" of its cut is quite unrealistic. Were the versions already ordered, the lowest version is surely on the low side of its cut, while the highest version is on the high side of its. We should be prepared to turn some of the cuts around.

Exhibit 2 starts by doing just this. We find, in panel A, that we can make all but one of the 12 ff-logs positive. by an easy choice of which to turn around. (As that one is of size .01, its sign is highly negligible.) Panel B gives a

exhibit 1 of chapter 30x -- hair and eyes

The Badeners in ff-logs
(see exhibit 2 of chapter 30)

A) The TAGS

Fair hair	1	Blue eyes	5
Brown hair	2	Green or Gray	6
Black hair	3	Brown eyes	7
Red hair	4		

B) The 4-by-3 ARRAY of 2-by-2 TABLES

Eye color	Hair color cut			
	1/not 1	2/not 2	3/not 3	4/not 4
5	1768 | 1043	807 | 2004	189 | 2622	47 | 2764
not 5	1061 | 2928	1825 | 2164	1034 | 2955	69 | 3920
6	946 | 2186	1387 | 1745	796 | 2386	53 | 3079
not 6	1883 | 1985	1245 | 2423	477 | 3191	63 | 3605
7	115 | 742	438 | 419	288 | 569	16 | 841
not 7	2714 | 3229	2194 | 3749	935 | 5008	100 | 5843

Examples: (see panel A of exhibit 1):

1768 = 1768

1001 = 946 + 115

1043 = 807 + 189 + 47

2928 = 1387 + 638 + 746 + 288 + 53 + 16

C) The ff-LOGS -- matched

	1/not 1	2/not 2	3/not 3	4/not 4
5/not 5	.77	-.37	-.79	-.02
6/not 6	-.39	.22	.37	-.01
7/not 7	-.84	.29	.50	.05

exhibit 1 (cont'd.)

D) An ELEMENTARY ANALYSIS by median differences, unit = .01

	1/	2/	3/	4/	eff	fit
5/	75	-26	-83	32	-22	-28
6/	3	3	3	3	8	2
7/	-49	4	10	3	14	8
eff	-43	17	32	-6	-6	0
fit	-49	23	38	12	0	6

Hinges of residuals are -11 and +7. Clarity = 111/18 = 6.

Commonality = 6/18 = 0.3.

E) PROBLEMS

1) Make an extended elementary analysis of panel C.

2) Plot either the residuals of panel D or those of panel E (or both separately) against the residuals of the elementary analysis of panel C of exhibit 3 of chapter 30.

exhibit 2 of chapter 30x -- hair to eye color

The ff-logs after turning some cuts around

A) The UNTWISTED ff-LOGS -- corresponding to reoriented cuts

	1/not 1	not 2/2	not 3/3	not 4/4
5/not 5	.77	.37	.79	.02
not 6/6	.39	.22	.37	-.01
not 7/7	1.98	.29	.50	.05

B) An ELEMENTARY ANALYSIS -- after 3 half-steps of medians, unit = 0.01

					eff	fit
	0	1	27	-5	2	44
	-24	0	-1	6	-12	30
	123	-5	0	0	0	42
eff	33	-8	8	-37	42	00
fit	75	34	50	5	00	-42

Hinges: -3 and +3h, Clarity = $\frac{70+14}{6h}$ = 13. Commonality = $\frac{42}{6h}$ = 6.5.

C) The ff-LOGS corresponding to the ORDERINGS suggested by panel B

	1/423	14/23*	142/3
5/67	.77	.76	.79
56/7	1.98	.78	.50

* See exhibit 4 for calculations, and comparison with 12/43.

D) The SMALLEST CORNERS corresponding to these 2-by-2 TABLES

5/67	1043	996	189
56/7	115	131	288

corresponding two-way analysis. Its clarity is 13, rather than 6 -- its commonality is 6.5 rather than 0.3. Clearly the combination of:

- which cuts are turned around,
- the elementary analysis,

tell us much more than did the elementary analysis of the previous exhibit (exhibit 1).

In particular, it tells us enough to do a good job of choosing orderings for the versions of each factor. To do this, we need to recognize what the effects in our elementary analysis mean. Large (positive) values place the corresponding version -- whether or not it has been turned around --

close to the end

of its order.

Clearly, then, 1/not 1 and not 3/3 are the extreme versions among 1,2,3,4. Also, not 2/2 is closer to its end than not 4/4. Hence the order to try is 1,4,2,3. For the other response, clearly 5/not 5 and not 7/7 are the extremes. This fixes the order as 5,6,7 (or the reverse, which would do as well)

Panel C shows the 6 ff-logs corresponding to the pairs of cuts consistent with this ordering. We see 4 of the 6 ff-logs between .76 and .79, with the other two straying, one in each direction. Clearly the two orderings we have selected,

fair hair, red hair, brown hair, black hair

and

blue eyes, green or grey eyes, brown eyes

seem to be quite plausible, when each is judged by its relation to the other. (They also seem to fit our preconceptions fairly well.)

* a weaker example -- Woo *

Exhibit 3 is based on Woo's 1928 counts on laterality of hand and eye, each classified as R = right, A = ambi-, or L = left. The nine 2-by-2 tables are shown in panel C. The ff-logs, shown in panel D, clearly point out L/AR and A/LR as the strong cuts for the eye -- and A/LR as the strong cut for the hand. As panel E makes clear, there is no evidence worthy of the name as to any order between right-handedness and left-handedness.

The results of the analysis are indications -- which might or might not be confirmed by more data -- that:

- as far as eyes go, we may wish to put left eyedness at one end and ambi-eyedness at the other.
- that ambidexterity may be different from both right and left-handedness.

The evidence here is weaker, but the technique applies.

exhibit 3 of chapter 30x -- hand and eye

Woo's Counts on Laterality of Hand and Eye

A) BASIC COUNTS

Hand*	Ocular laterality for general astigmatism		
	"Left-eyed"	Ambiocular	"Right-eyed"
Left-handed	34	62	28
Ambidextrous	27	28	20
Right-handed	57	105	52

*Manual laterality as measured by a balancing device.

B) SOURCE: T. L. Woo 1928. Dextrality and sinistrality of hand and eye. Second Memoir. Biometrika 20A: 79-148

C) 2-by-2 TABLES

Cut for hand	Cut for eye					
	L/AR		A/LR		R/AL	
L / AR	34	90	62	62	28	96
	84	205	133	156	72	217
A / LR	27	48	28	47	20	55
	91	247	167	171	80	256
R / AL	57	157	105	109	52	162
	61	138	90	109	48	151

exhibit 3 (cont'd.)

D) ff-LOGS

Cut for hand	Cut for eye		
	L/AR	A/LR	R/AL
L/AR	-.04	.08	-.06
A/LR	.21	-.24	.07
R/AL	-.10	.08	.00

Clear choices: -.24 and then .21, giving A-to-R-to-L order for eye (or the reverse) and A distinctive for hand.

E) R vs L for HAND

Hand	A	R	L	
	observed			
R	105	52	57	(214)
L	62	28	34	(124)
	fitted from marginal totals			
R	105.6	50.7	57.6	
L	61.4	29.3	33.4	
	residuals			
R	-.6	1.3	-.6	
L	+.0	-1.3	.6	

Clearly the relationship of right- and left-handedness to laterality of eyes essentially vanishes in this data.

30xC. "All cuts" approaches

So far we have been looking at only some of the possible 2-by-2 tables as a basis from which to choose our orders -- specifically those in which a single version labels one row and a single version labels one column. Doing this holds down the computation considerably. If all we need is a plausible order, looking at the smaller bundle is often enough. Sometimes, it is nice to go further.

* back to the Badeners *

Let us return to the hair and eye color of the Badeners, who provide a relatively simple case for which we have done much of the work. If a response has only three versions, any useful cut divides these one and two. No new cuts are to be added to those which isolate one of the versions. If a response has four versions, there are seven cuts that we need to consider -- four that divide the versions one and three, three more (not yet considered) that divide them two and two.

Exhibit 4 begins by calculating the ff-logs for the remaining 2-by-2 tables. Panel B shows the complete 3-by-7 table of ff-logs. Panel C shows the result of turning around those cuts that seem in need of this, and of further rearranging cuts in order of size -- as measured by the sum of the corresponding ff-logs.

exhibit 4 of chapter 30x -- hair and eye

The remaining 2-by-2 tables and an analysis
combining all 3 × 7 = 21
(see exhibit 1 for tags and sources)

A) The additional 3-by-3 in 2-by-2 TABLES

Eye color	Hair color cut 12/34		Hair color cut 13/24		Hair color cut 14/23	
5/not 5	2575	236	1957	854	1815	996
	2886	1103	2095	1894	1130	2859
6/not 6	2333	849	1742	1440	999	2183
	3128	540	2360	1308	2360	1308
7/not 7	553	304	403	454	131	726
	4908	1035	3649	2294	2814	3229

B) All 7-by-3 ff-LOGS -- matched

	1/234	2/134	3/124	4/123	12/34	13/24	14/23
5/not 5	.77	-.37	-.79	-.02	.71	.36	.76
6/not 6	-.39	.22	.37	-.01	-.37	-.20	-.69
7/not 7	-1.98	.29	.50	.05	-.49	-.29	-.78

C) Some CUTS INVERTED, all CUTS REARRANGED by SIZE -- unit = 0.01

	1/234	14/23	124/3	12/34	134/2	13/24	123/4
5/67	77	76	79	71	37	36	.02
56/7	198	78	50	49	29	29	.05
57/6	39	69	37	37	22	20	-.01
(sum)	(314)	(223)	(166)	(157)	(88)	(85)	(6)

If all we wanted was orderings, panel C would give us the same orderings we found in the last section. Panel C does, however, offer us something more. If we look hard at the order of cuts among 1,2,3, and 4, we see that -- starting from large sizes -- we have three successive pairs each of which differs only in where version 4 is placed. The seventh cut involves 4 alone and has a very small sum. The reasonable diagnosis from all this is that version 4 (red hair), either because it occurs rather infrequently or because it is really not part of the order, does not have very much influence on the ff-logs. (The obvious exception involves 1/234 and 14/23 with 56/7 and 57/6. Clearly some further anomalies seem to be occurring among these four 2-by-2's, so we can probably live with the generalization.)

For one thing, this leads us to look at the ff-logs we get if we drop version 4 entirely. Exhibit 5 has the ff-logs, their two-way analysis, and the resulting orderings. (We are not surprised to see our old familiar friends.)

Another thing we can do is to try to fit version 4 into the 1,2,3 order more carefully. A possibility is to look at only the extreme versions of eye color. Exhibit 6 carries out the arithmetic. Red hair is placed in the same order, confirming our previous analysis -- of rather more interest is the scaling of the four hair colors.

exhibit 5 of chapter 30x -- hair and eye

The result of dropping version 4 (red hair)
(for data and logs see exhibit 1 of chapter 30)

A) The ff-LOGS -- matched

	1/23	12/3	13/2
5/67	.78	.79	.37
56/7	.83	.50	.30
57/6	.60	.37	.22

B) A TWO-WAY ANALYSIS -- unit = 0.01 in matched ff-logs, three half-steps of median polish

			fit
0	17	0	7
12	-5	0	0
-3	0	6	-8
16	0	-30	-55

C) RESULTING ORDERINGS

1/23 and 12/3, hence 1,2,3.

5/67 and 56/7, hence 5,6,7.

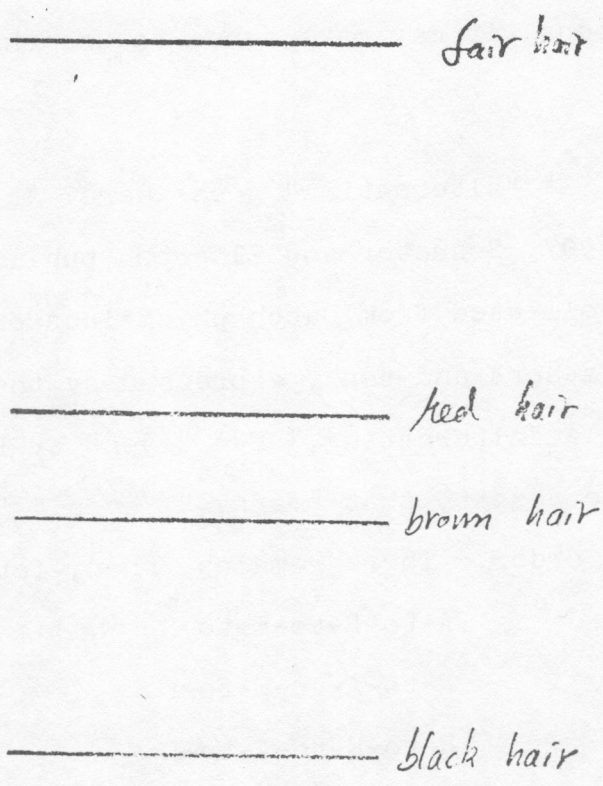
exhibit 6 of chapter 30x -- hair and eye

Fitting version 4 -- red hair -- into the order of the other 3 hair colors by its relationship to the extremes of eye color

A) DATA and RESULTS

Eye Color	Hair color 1	2	3	4	
5	1768	807	189	47	counts
7	115	438	288	16	
%5	93.9	64.8	39.6	74.6	%
f + log	1.36	.31	-.21	.54	

B) PICTORIAL VIEW of ORDER and APPARENT SPACING -- as shown by f-logs



* unrestricted or restricted? *

We must beware of trying "all cuts" analysis, at least by hand, with many versions. The number of cuts needed for each response runs as follows:

three versions----------3 cuts

four versions----------7 cuts

five versions----------15 cuts

six versions----------31 cuts

Thus doing "all cuts" on a 6-by-6 would require 31 × 31 = 961 different 2-by-2 tables.

30xD. <u>A restricted example</u>

In many instances, however, there are plausible arguments for looking at only a limited number of possibilities. When this happens, it may make sense to use this approach by hand.

* "alternating" goes where? *

In 1907, Schuster and Elderton published an analysis of some data collected from Dutch physicians on the general character of fathers and sons, expressed by their being R = "merry", A = "alternating," E = "even" or L = "melancholy." It is clear, we assert, that "merry," "even," "melancholy" belong in that order. There remains, then, four possible orders:

A-to-R-to-E-to-L

R-to-A-to-E-to-L

R-to-E-to-A-to-L

R-to-E-to-L-to-A

Whichever one we accept ought to apply to both fathers and sons. Of the 7 cuts that are possible with 4 classes, only two -- LR/AE and E/ALR -- are excluded. Of the 5 × 5 = 25 combinations remaining, two -- RA/EL with RE/AL and RE/AL with RA/EL -- are excluded. The other 23 are possible, sometimes with a reversal of sign.

Exhibit 7 gives the counts, the 23 2-by-2 tables, and the corresponding ff-logs. Nine of the 23 2-by-2 tables involve one very small cell. If we set these aside -- which guarantees that L will not be an extreme class -- we get the order of sizes of ff-logs shown in panel E, which leads to the order

Merry-to-even-to-melancholy-to-alternating

and to the elementary analysis of panel G, which is mainly the common term = .21 with some tendency for the effects for R/AEL to be a few hundredths positive.

To see what we lost by omitting ff-logs with one small cell, panels H to J repeat the analysis keeping all ff-logs. The resulting order is now

Merry-to-alternating-to-even-to-melancholy

with an elementary analysis with a smaller common term, larger effects, and two quite large residuals.

This second choice of order depended on an ff-log of .39 involving a cell with a count of 2. Had this been 3

exhibit 7 of chapter 30x -- psychical inheritance

Inheritance of psychical characters

A) The COUNTS

Son	Father: R merry	A alternating	E even	L melancholy	(sum)
R merry	122	81	67	8	(278)
A alternating	70	101	68	9	(248)
E even	58	66	45	6	(175)
L melancholy	10	7	10	2	(29)
(sum)	(260)	(255)	(190)	(25)	730

B) SOURCE

E. Schuster and E. M. Elderton 1907. The inheritance of psychical characters. Biometrika 5 = 460-469. Table XV(a) on page 466.

Their source: G. Heymans and E. Wiersma 1906. Beiträge zur speziellen Psychologie auf Grund einer Massenuntersuchung. Zeitschr. für Psychologie 42:1-127 and 258-301.

C) The DOUBLE DICHOTOMIES

	fathers: R/AEL		RA/EL		RAE/L		RE/AL		REL/A	
R / AEL	122	156	203	75	270	8	189	89	197	81
	138	314	312	140	435	17	261	191	278	164
RA / EL	192	334	374	152	509	17	XXXXXXXX		345	181
	68	136	141	63	196	8	XXXXXXXX		133	73
RAE / L	250	651	498	203	678	23	430	271	453	248
	10	19	17	12	27	2	20	9	22	7
RE / AL	180	273	XXXXXXXX		439	14	292	161	306	147
	80	197	XXXXXXXX		266	11	158	119	169	108
REL / A	190	292	344	138	239	9	312	170	328	159
	70	178	171	77	466	16	138	110	147	101

(Cont'd)

exhibit 7 (cont'd.)

D) The ff-LOGS -- with count of smallest cell in ( )

son	R/AEL	RA/EL	RAE/L	RE/AL	REL/A
R/AEL	.28 (122)	.10 (75)	.14 (8)	.22 (89)	.18 (81)
RA/EL	.07 (68)	.05 (63)	.10 (8)	XXXXX	.02 (73)
RAE/L	-.16 (10)	.27 (12)	.39 (2)	-.17 (9)	-.27 (7)
RE/AL	.24 (80)	XXXXX	.13 (11)	.15 (119)	.14 (108)
REL/A	.24 (70)	.06 (77)	-.04 (9)	.19 (110)	.19 (101)

E) ORDERED ARRANGEMENT of SIZE of ff-LOGS -- for at least 50 in the smallest cell

size of ff-log	Father	son	selection** 1st R/	2nd RE/	3rd REL/
.28*	R/	R/	✓	✓	✓
.24	R/	RE/		✓	✓
.24	R/	REL/			✓
.22	RE/	R/		✓	✓
.19*	RE/	REL/			✓
.19*	REL/	REL/			✓
.18	REL/	R/			✓
.15*	RE/	RE/		✓	✓
.14*	REL/	RE/			✓
.10	RA/	R/			
.07	R/	RA/			
.06	RA/	REL/			
.05	RA/	RA/			
.02	REL/	RA/			

* smallest cell at least 100.

** check when dichotomies included in this column or in the column(s) to left.

exhibit 7 cont'd.

Note: Giving only the left-hand side of the cut makes looking at the table much easier, as the reader may verify by copying this panel with full labels, such as R/AEL, RE/AL, etc.

F) SELECTED ff-LOGS

	Father		
Son	R/AEL	RE/AL	REL/A
R/AEL	.28	.22	.18
RE/AL	.24	.15	.14
REL/A	.24	.19	.19

G) ELEMENTARY ANALYSIS -- unit = 0.01

	R/AEL	RE/AL	REL/A	eff	fit
R/AEL	0	1	-2	2	23
RE/AL	1	-1	-1	-3	18
REL/A	-2	0	1	0	21
eff	5	-2	-3	21	00
fit	26	19	18	00	=21

Hinges: -1 and +1 Centrality: $10h = \frac{21}{2}$

Clarity: $6h = \frac{13}{2}$. Outs: 0/9

(Cont'd)

exhibit 7 (cont'd.)

H) ORDERED ARRANGEMENT SIZES ff-LOGS -- all legal double dichotomies, no matter how small the count.

Size of ff-log	Father	Son	Smallest cell	Selection 1st RAE/	2nd R/	3rd RA/
.39	RA/	RAE/	(2)	✓		
.28	R/	R/	(122)		✓	
.27	RA/	RAE/	(12)			✓
.27	REL/	RAE/	(7)			
.24	R/	RE/	(80)			
.24	R/	REL/	(70)			
.22	RE/	R/	(89)			
.19	RE/	REL/	(110)			
.19	REL/	REL/	(101)			
.18	REL/	R/	(81)			
.17	RE/	RAE/	(9)			
.16	R/	RAE/	(10)		✓	
.15	RE/	RE/	(119)			
.14	REL/	RE/	(108)			
.14	RAF/	R/	(8)		✓	
.13	RAE/	RE/	(11)			
.10	RA/	R/	(25)			✓
.10	RAE/	RA/	(8)			✓
.07	R/	RA/	(68)		✓	
.06	RA/	REL/	(77)			
.05	RA/	RA/	(63)		✓	
.04	RAE/	REL/	(9)			
.02	REL/	RA/	(73)			

(Cont'd)

exhibit 7 (cont'd.)

I) SELECTED ff-LOGS

Son	Father		
	R/AEL	RA/EL	RAE/L
R/AEL	.39	.10	.14
RA/EL	.07	.05	.10
RAE/L	.16	.27	.28

selected order: R-to-A-to-E-to-L.

J) ELEMENTARY ANALYSIS -- unit = 0.01

	R/AEL	RA/EL	RAE/L	eff	fit
R/AEL	27	0	0	-3	12
RA/EL	0	0	1	-8	7
RAE/L	-8	0	1	10	25
eff	0	-2	2	15	00
fit	15	13	17	00	-15

K) PROBLEM

1) Repeat the selection of an order, using the values of

$$(\text{size of ff-log}) \times \sqrt{\text{count in smallest cell}}$$

as the basis for selection.

2) Try the approach of the last section, using only cuts with a single version on one side.

or 1, the ff-log would have been .18 or .74. This is a rather insecure base for an analysis.

The choice left to the reader's efforts in panel K, using

$$(\text{size of ff-log})\sqrt{\text{smallest count}}$$

as the guide is probably the approach to be recommended.

****

30xH. How far have we come?

Even if a set of versions is not ordered, those cuts that divide one version from all others have a special position. If we consider all such for both factors, a two-way table is converted into a two-way bundle of 2-by-2 tables. Each bundle gives us an ff-log. If we are willing to turn any one cut around, all its ff-logs change sign. By careful choice of sign changing we may be able to make the ff-logs predominantly positive.

Once this is done, a two-way analysis of these partly-turned-around ff-logs can tell us which versions are at the ends of a reasonable ordering -- and then which come next. This process often orders both factors completely, though weaker evidence -- leading to incomplete ordering -- is not infrequent.

If there are not too many versions, say four, we can also try all possible cuts -- including those leaving two or more versions on either side.

When we have other information we can try less than all cuts but not just those that separate off a single version.

****

30xP. <u>Problems for chapter 30x</u>

1. Do problem 1 of exhibit 1.
2. Do problem 2 of exhibit 1.
3. Do problem 1 of exhibit 7.
4. Do problem 2 of exhibit 7.